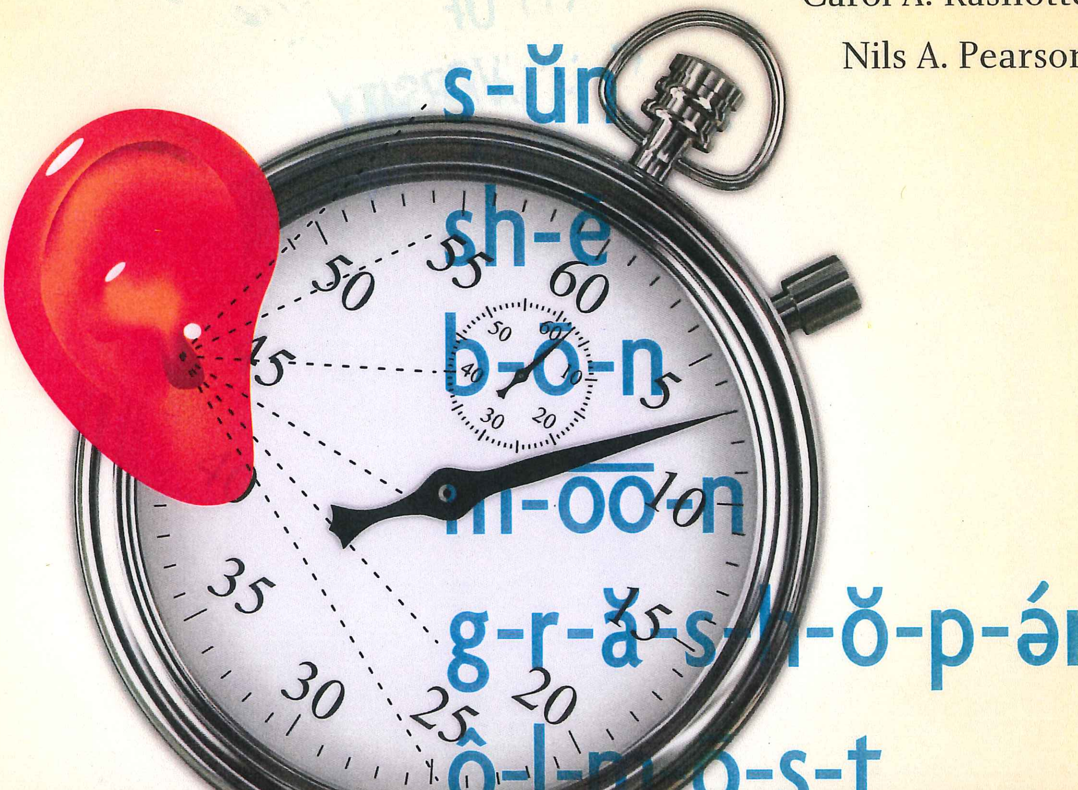
CTOPP 2

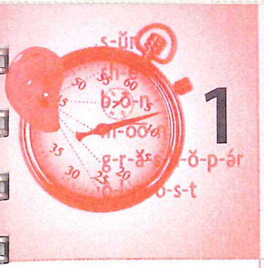
Comprehensive Test of Phonological Processing

Second Edition

Examiner's Manual

Richard K. Wagner
Joseph K. Torgesen
Carol A. Rashotte
Nils A. Pearson





1 Introduction to the CTOPP-2

The CTOPP-2 is a norm-referenced test that measures phonological processing abilities related to reading. The methods used to build the CTOPP-2 and the procedures for giving, scoring, and interpreting it are described later in this manual, along with evidence for its reliability and validity. Before dealing with these topics, however, we provide some introductory information about phonological processing and the CTOPP-2. This chapter describes (a) the nature of phonological processing and its relations to speech and reading; (b) the test model used to build the CTOPP-2; (c) the CTOPP-2's components, subtests, and composite scores; and (d) uses of the test.

The Nature of Phonological Processing and Its Relations to Speech and Reading

According to Crystal (2001) and other linguists, the term *phonology* refers to the sound system of language. The most important component of phonology, often called *phonemics*, involves the study of significant speech sounds. The sounds (phonemes) are the building blocks we use to construct words and sentences when speaking. Individuals who have problems processing phonological material may mispronounce words when speaking or misperceive the speech of others. They may also experience difficulty in other speech abilities, such as rhyming, segmenting sounds in words, blending phonemes, and discriminating between speech sounds.

Traditionally, interest in phonology has been held mostly by linguists, speech pathologists, and speech scientists. Over the past 40 years, however, many psychologists and educators have discovered that some phonological processing abilities play an important role in learning to read and write. Evidence that some phonological processing abilities are involved in learning to read and write comes from both correlational and training studies. Three kinds of phonological processing in particular appear to be especially relevant to the development of written language: phonological awareness, phonological memory, and rapid naming.

Phonological Awareness

Phonological awareness refers to an individual's awareness of and access to the sound structure of his or her oral language (Mattingly, 1972). The spoken words of a language represent strings of phonemes that signal differences of meaning. The spoken word *cat* has three phonemes, each of which happens to correspond to the sound made by the three letters of the printed word *CAT*. Change the first

sound from /k/ to /b/ and you have the spoken word *bat*. Of the nearly infinite number of possible strings of phonemes that might be represented in an oral language, only a relatively small number actually occur, and these occur in multiple words. Children who have some awareness of this structure seem to have an advantage in learning to read the printed forms of a language. This makes sense given that many printed languages attempt to convey pronunciation as well as meaning, and this is even more true for alphabetic languages such as printed English because letters of the alphabet have a rough correspondence to phonemes.

As children develop, they demonstrate awareness of increasingly smaller phonological units of speech. Initially, their awareness is limited to word-length phonological units, as in recognizing the two parts of the compound word *cowboy*. Next, they become aware of syllables within words, as in recognizing each syllable of the two-syllable word *seven*. Eventually, awareness proceeds within the syllable to recognition of onsets and rimes. For example, for the first syllable of the word *seven*, the onset is the sound of the initial *s*, and the rime is the sound of the vowel and remaining consonant *ev*. Eventually, awareness is demonstrated for individual phonemes within syllables. Awareness occurs for individual phonemes within rimes, as when recognizing the sounds represented by the letters *e* and *m* in the word *stem*. Subsequently, awareness of individual phonemes within consonant clusters occurs, as when recognizing the sounds represented by the letters *s* and *t* in *stem*.

An important reason for assessing phonological awareness is that clear implications for intervention exist. Many children who are weak in phonological awareness show improvement in their word-level reading skills after being given intervention designed to improve their phonological awareness (Ball & Blachman, 1991; Bradley & Bryant, 1985; Brady, Fowler, Stone, & Winbury, 1994; Cunningham, 1990; Ehri et al., 2001; Lundberg, Frost, & Petersen, 1988; National Reading Panel, 2000). In general, reading approaches that feature systematic, explicit instruction in phonological awareness and phonemic decoding skills produce stronger reading growth in children who are weak in phonological awareness compared with reading approaches that do not teach these skills explicitly (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Hatcher, Hulme, & Ellis, 1994; Lovett, Borden, Lacerenza, Benson, & Brackstone, 1994; Scammacca, Vaughn, Roberts, Wanzek, & Torgesen, 2007; Torgesen, Wagner, & Rashotte, 1997; Torgesen, Wagner, et al., 1999). Nearly all reading series that are used to teach young children to read target developing phonological awareness, a practice that was supported by a review for the What Works Clearinghouse Practice Guide on Foundational Reading Skills (Foorman et al., 2012).

Phonological Memory

Phonological memory refers to coding information phonologically for temporary storage in working or short-term memory. When you attempt to remember a phone number you have looked up as you make your way to the phone, you are storing the number temporarily in working memory. You probably do so not by storing a visual representation of the sequence of digits (although you may be

able to do this if you try) but rather by storing a phonological representation of the sounds of the digit names.

The part of memory most involved in storing phonological information is called the *phonological loop*. The phonological loop provides a brief, verbatim storage of auditory information (Baddeley, 1986, 1992; Torgesen, 1996). The phonological loop consists of two parts working together. The first is a phonological store, which can be thought of as a tape-recording loop that retains the most recent 2 seconds' worth of auditory information that has been recorded. The second is an articulatory control process that provides input to the phonological loop initially and also can refresh information already in the loop so that it can be stored for longer than 2 seconds.

A deficient phonological memory does not appear to impair either word-level reading or listening to a noticeable extent, providing that the words involved are already in the individual's vocabulary. However, phonological memory impairments can constrain the ability to learn new written and spoken vocabulary (Gathercole & Baddeley, 1990; Gathercole, Willis, & Baddeley, 1991). One clear example is provided by a study of an individual who, as an adult, suffered a brain injury that selectively impaired phonological memory (Baddeley, Papagno, & Vallar, 1988). This individual could converse and read relatively fluently provided the words were in his spoken and reading vocabularies. However, this native English speaker was unable to learn new words in the form of Spanish vocabulary. In fact, he was unable to even repeat Spanish words accurately. In addition, when presented orally with brief, two-syllable nonwords, his accuracy for repeating them was less than 50%. Studies of elementary school–age children also indicate that deficits in phonological memory result in impaired development of written and spoken vocabularies (Gathercole & Baddeley, 1990).

Perhaps the most comprehensive investigation of impaired phonological memory is a series of more than 20 experiments that investigated the consequences of poor phonological memories for children with reading disabilities (Torgesen, 1988, 1996; Torgesen & Houck, 1980; Torgesen, Rashotte, & Greenstein, 1988). Early studies demonstrated that the origin of the memory deficit was a specific deficit in phonological coding of familiar verbal materials, such as digits and words. Alternative explanations, such as inattention, anxiety, poor use of mnemonic strategies, and poor motivation, were ruled out. The children were not impaired in short-term memory for nonverbal material, long-term memory, or listening comprehension; however, they showed severe impairments in decoding visually presented nonsense syllables. When followed up with approximately a decade later, more than half of the individuals still had severe memory impairments, and those who did showed almost no improvement in reading skills. Finally, consistent with the studies of Gathercole and Baddeley (1990), individuals with impaired phonological memories as children scored poorly on the Vocabulary subtest of the *Wechsler Adult Intelligence Scale–Revised* (Wechsler, 1981) a decade later, yet performed adequately on nonverbal performance subtests such as Block Design.

Phonological coding in working memory may not be all that important for reading common known words, although some controversy exists about this assertion (Crowder & Wagner, 1991). It is clear, however, that phonological coding in working memory is potentially more useful when attempting to decode new

words, particularly words that are long enough to decode bit by bit, as a means of storing intermediate sounds.

Phonological processing also appears to be important to reading by supporting the role of working memory in comprehension (Rayner, Pollatsek, Ashby, & Clifton, 2011). The meaning of a sentence depends not only on the words it contains but also on their order in the sentence. "The cat chased the dog" differs in meaning from "the dog chased the cat," even though the words are identical.

Rapid Naming

The third kind of phonological processing is rapid naming. Rapid naming of digits, letters, objects, or colors requires efficient retrieval of phonological information from long-term, or permanent, memory. Unlike phonological awareness and phonological memory (which are entirely auditory–oral in mode), rapid naming has visual components, most of which are graphemes or glyphs. Because of this, rapid naming is best thought of as being a hybrid ability, in that successful performances depend on how fast an examinee can scan the array of visual symbols and encode a phonological response. Obviously, the mixed modality nature of this ability is the same type that underlies decoding when reading aloud.

When reading, young readers presumably retrieve (a) phonemes associated with letters or letter pairs, (b) pronunciations of common word segments, and (c) pronunciations of whole words. The efficiency with which children are able to retrieve phonological codes associated with individual phonemes, word segments, or entire words should influence the degree to which phonological information is useful in decoding printed words (Baddeley, 1986; Wolf, 1991). The efficiency with which phonological codes are associated with items to be named will be affected by how well the items are known and how strong the mapping is between the item and its pronunciation.

Measures of rapid naming require speed and processing of visual as well as phonological information. Some researchers who study rapid naming suggest that rapid naming tasks assess the operation of a precise timing mechanism that is important for the developing knowledge of common letter patterns in printed words (Bowers & Wolf, 1993; Wolf, 1991). Others have focused on the visual aspects of the rapid naming task as the reason it is predictive of reading and problematic for individuals with reading disabilities (Geiger, Lettvin, & Zegarra-Moran, 2009; Stein & Walsh, 1997), or have hypothesized that poor performance on rapid naming tasks is due to domain-general problems in speed of processing (Kail & Hall, 1994; Kail, Hall, & Caskey, 1999). Subsequent studies have not supported these alternative hypotheses, however (Bonifacci & Snowling, 2008; Hawelka & Wimmer, 2005; Lervag & Hulme, 2009).

It is clear that rapid naming tasks predict poor performance in reading, and they do so independently of measures of phonological awareness (Bowers & Swanson, 1991; Lervag & Hulme, 2009; Manis, Doi, & Badha, 2000; Parrila, Kirby, & McQuarrie, 2004; Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007). Individuals who have double deficits—that is, deficits in both rapid naming and phonological awareness—appear to have greater difficulty learning to read words accurately and fluently than do individuals with deficits in either rapid naming or phonological awareness alone (Bowers & Wolf, 1993). What is

less clear is exactly what rapid naming tasks measure that is not captured by measures of phonological awareness or phonological memory. The most promising candidate appears to be something related to the ability to connect visual stimuli to phonological codes, as opposed to the adequacy of the phonological codes by themselves (Jones, Branigan, Hatzidaki, & Obregon, 2010; Manis, Seidenberg, & Doi, 1999; Wimmer, Mayringer, & Landerl, 1998).

A Model of Phonological Processing

The frame of reference described in the previous section can be depicted in a model of reading-related phonological processing (Wagner et al., 1987; Wagner & McBride-Chang, 1996; Wagner & Torgesen, 1987; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner, Torgesen, & Rashotte, 1994; Wagner et al., 1997). This model is presented in Figure 1.1.

Phonological awareness, phonological memory, and rapid naming represent three correlated yet distinct kinds of phonological processing abilities. These abilities are correlated rather than independent in that confirmatory factor analytic studies reveal that the correlations between them are substantially greater than 0. They are distinct rather than undifferentiated in that the correlations between them are less than 1. In general, phonological awareness and phonological memory tend to be more highly correlated with one another than with rapid naming. In addition, the three kinds of phonological processing abilities tend to become less correlated with development. For very young children, phonological awareness and phonological memory can be nearly perfectly correlated (Wagner et al., 1987; Wagner et al., 1993).

Description of the CTOPP-2

The model previously described is the theoretical foundation underlying the development of the three constructs representing phonological abilities: phonological awareness, phonological memory, and rapid naming. The CTOPP-2 was developed to aid in the identification of individuals from kindergarten through college who may profit from instructional activities to enhance their phonological skills.

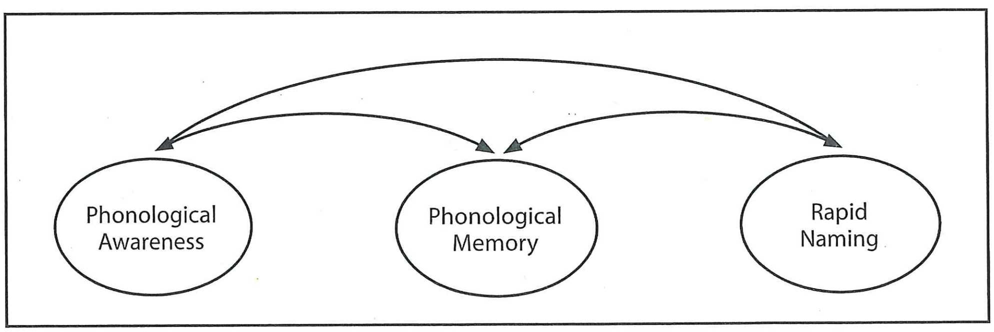


Figure 1.1. The model of phonological processing.

Description of the Test Components

The CTOPP-2 test kit contains an Examiner's Manual, a Picture Book, an Examiner Record Booklet for ages 4–6, an Examiner Record Booklet for ages 7–24, and two audio CDs. One CD is used to present the core subtests Blending Words, Memory for Digits, and Nonword Repetition, and the other CD is used to present the supplemental subtests Blending Nonwords and Segmenting Nonwords.

Description of the Subtests

The model described in the previous section contained three reading-related phonological processing abilities: phonological awareness, phonological memory, and rapid naming. The CTOPP-2 subtests were designed to assess skills within each of the three constructs. The core and supplemental subtests listed below are based on experimental tasks that have been used to study phonological processing in the published literature. We revised and evaluated these tasks extensively in a decade of research (Wagner et al., 1987; Wagner & Torgesen, 1987; Wagner et al., 1993; Wagner et al., 1994; Wagner et al., 1997).

Elision (Core, 4–24 Years).

This 34-item subtest measures the extent to which an individual can say a word and then say what is left after dropping out designated sounds. For the first two items, the examiner says compound words and asks the examinee to say that word and then say the word that remains after dropping one of the compound words. For the remaining items, the individual listens to a word and repeats that word and then is asked to say the word without a specific sound. For example, the examinee is instructed, "Say 'bold.'" After repeating "bold," the examinee is told, "Now say 'bold' without saying /b/." The correct response is "old."

Blending Words (Core, 4–24 Years).

This 33-item subtest measures an individual's ability to combine sounds to form words. The examinee listens to a series of audio-recorded separate sounds and then is asked to put the separate sounds together to make a whole word. For example, the examinee is asked, "What word do these sounds make: t-oi?" The correct response is "toy."

Sound Matching (Core, 4–6 Years).

This 26-item subtest measures the extent to which an individual can match sounds. The examiner says a word, pauses, and then says three other words while pointing to drawings depicting all four words. The pictures in the Picture Book are used to help the examinee remember all of the words. For the first 13 items, the examinee is asked to point to the picture that corresponds to the word that starts with the same sound as the word the examiner provides. For the last 13 items, the same procedure is used, except the individual is asked to point to the picture of the word that ends with the same last sound as the word the

examiner provided. For example, the examiner asks the examinee, "Which word starts with the same sound as *pan*? *Pig*, *hat*, or *cone*?" The correct response is "pig."

Phoneme Isolation (Core, 7–24 Years).

This 32-item subtest measures an individual's ability to identify target sounds in words. The items begin with CVC words, and the task is to identify the first and then last sounds of presented words. Subsequent items require the examinee to identify middle sounds and then proceed to words made up of more sounds. The most difficult items require identifying a named sound in words that have more letters than sounds, for example, "What is the *second* sound in the word *island*?" These items are difficult because the correct answer cannot be obtained by using a spelling strategy of simply naming the sound of the second letter in the word.

Memory for Digits (Core, 4–24 Years).

This 28-item subtest measures the extent to which an individual can repeat a series of numbers ranging in length from two to eight digits. After the individual has listened to a series of audio-recorded numbers presented at a rate of 2 per second, he or she is asked to repeat the numbers in the same order in which they were heard.

Nonword Repetition (Core, 4–24 Years).

This 30-item subtest measures an individual's ability to repeat nonwords that range in length from 3 to 15 sounds. The examinee is told to listen to an audio-recorded made-up word and repeat it exactly as he or she heard it. For example, the examinee hears the recorded sounds "nigong," to which the correct response is "nigong."

Rapid Digit Naming (Core, 4–24 Years).

This 36-item subtest measures the speed with which an individual can name numbers. The Picture Book contains one page for this subtest, which consists of four rows and nine columns of six randomly arranged numbers (i.e., 2, 3, 4, 5, 7, 8). The examinee is instructed to name the numbers on the top row from left to right, and then name the numbers on the next row from left to right, and so on, until all of the numbers have been named. The individual's score is the total number of seconds taken to name all of the numbers on the page.

Rapid Letter Naming (Core, 4–24 Years).

This 36-item subtest measures the speed with which an individual can name letters. The Picture Book contains one page for this subtest, which consists of four rows and nine columns of six randomly arranged letters (i.e., *a*, *c*, *k*, *n*, *s*, *t*). The examinee is instructed to name the letters on the top row from left to right, and then name the letters on the next row from left to right, and so on, until all of the

letters have been named. The individual's score is the total number of seconds taken to name all of the letters.

Rapid Color Naming (Core, 4–6 Years).

This 36-item subtest measures the speed with which an individual can name the colors of a series of different colored blocks. The Picture Book contains one page for this subtest, which consists of four rows and nine columns of six randomly arranged colors (i.e., blue, red, green, black, yellow, brown). The individual is told to name the colors on the page as quickly as possible. He or she is instructed to name the colors on the top row from left to right and then name the colors on the next row from left to right, and so on, until all of the colors have been named. The individual's score is the total number of seconds taken to name all of the colors.

Rapid Object Naming (Core, 4–6 Years).

This 36-item subtest measures the speed with which an individual can name a series of objects. The Picture Book contains one page for this subtest, which consists of four rows and nine columns of six randomly arranged drawings of objects (i.e., pencil, star, fish, chair, boat, key). The examinee is instructed to name the objects on the top row from left to right and then name the objects on the next row from left to right, and so on, until all of the objects have been named. The individual's score is the total number of seconds taken to name all of the objects.

Blending Nonwords (Supplemental, 4–24 Years).

This 30-item subtest measures an individual's ability to combine speech sounds to make nonwords. The individual listens to a series of audio-recorded separate sounds and is then asked to put these separate sounds together to form a nonword. For example, the examinee is asked, "What made-up word do these sounds make: nim-by?" The correct response is "nimby."

Segmenting Nonwords (Supplemental, 7–24 Years).

This 31-item subtest measures an individual's ability to say the separate phonemes that make up a nonword. The examinee listens to an audio-recorded nonword, repeats the nonword, and then says it one sound at a time. For example, the examinee listens to the audio-recorded sounds "ren," repeats the nonword, and then says the nonword one sound at a time. The correct response is "r-ĕ-n."

Description of the Composites

In addition to the subtests, the CTOPP-2 has five composite scores. The scaled scores from two or more subtests are summed and then converted to composite scores, as shown in Table 1.1.

Table 1.1
CTOPP-2 Composites and Subtests

	Composites 4–6-year-olds				Composites 7–24-year-olds			
Subtest	Phonological Awareness	Phonological Memory	Rapid Symbolic Naming	Rapid Non-Symbolic Naming	Phonological Awareness	Phonological Memory	Rapid Symbolic Naming	Alternate Phonological Awareness
Core								
Elision	X				X			
Blending Words	X				X			
Sound Matching	X							
Phoneme Isolation					X			
Memory for Digits		X				X		
Nonword Repetition		X				X		
Rapid Digit Naming			X				X	
Rapid Letter Naming			X				X	
Rapid Color Naming				X				
Rapid Object Naming				X				
Supplemental								
Blending Nonwords								X
Segmenting Nonwords								X

Phonological Awareness Composite Score (PACS).

For 4–6-year-olds, Elision, Blending Words, and Sound Matching are the core subtests that are used to form the Phonological Awareness composite. For 7–24-year-olds, Elision, Blending Words, and Phoneme Isolation are the core subtests used to form the PACS.

Phonological Memory Composite Score (PMCS).

For all ages, Memory for Digits and Nonword Repetition are the core subtests that form the PMCS.

Rapid Symbolic Naming Composite Score (RSNCS).

For 4–24-year-olds, Rapid Digit Naming and Rapid Letter Naming are the core subtests for the RSNCS.

Rapid Non-Symbolic Naming Composite Score (RNNCS).

For 4–6-year-olds, Rapid Color Naming and Rapid Object Naming are the core subtests for the RNNCS. These subtests are not to be use with older individuals.

Alternate Phonological Awareness Composite Score (APACS).

For 7–24-year-olds, an alternate phonological awareness composite score can be formed by adding the scaled scores from the Blending Nonwords and Segmenting Nonwords subtests. This quotient is not available for individuals 4–6 years old.

Uses of the CTOPP-2

The CTOPP-2 has four principal uses: to identify individuals who are significantly below their peers in important phonological abilities, to determine strengths and weaknesses among developed phonological processes, to document individuals' progress in phonological processing as a consequence of special intervention programs, and to serve as a measurement device in research studies investigating phonological processing.

First, the CTOPP-2 can be used to obtain information that is helpful in identifying those children and young adults who are markedly deficient in phonological processing ability. As described previously, phonological processing abilities have been shown to be related to later success in acquiring beginning reading skills. In addition, a deficit in one or more aspects of phonological processing is viewed as the primary cause of the majority of cases of learning disabilities (Ackerman & Dykman, 1993; Bruck, 1990; Bruck & Treiman, 1990; Catts, 1989; Elbro, Nielsen, & Petersen, 1994; Felton & Brown, 1990; Felton & Wood, 1990; Hulme & Snowling, 2009; McBride-Chang, 1995a; McBride-Chang & Manis, 1996; Olson, Wise, Conners, Rack, & Fulker, 1989; Shankweiler & Liberman, 1989; Siegel & Ryan, 1988; Stanovich, 1988; Torgesen, 1991; Torgesen, Wagner, Simmons, & Laughon, 1990; Tunmer, 1989; Wagner, 1986; Wolf & Obregon, 1992). The CTOPP-2 can be used to verify or document a processing deficit that is suspected on the basis of an individual's performance on other measures, such as the *Wechsler Intelligence Scale for Children–Fourth Edition* (Wechsler, 2003). For this usage, one might administer subtests corresponding to one, two, or all three kinds of phonological processing, depending on one's purpose. A particularly valuable use of the CTOPP-2 in this context is early identification of kindergarten and first-grade children whose deficiencies in phonological processing leave them at risk for reading failure. Early identification coupled with early intervention designed to promote phonological awareness appears to be a promising approach to reducing the incidence of reading failure.

Second, the results of the CTOPP-2 are useful to inventory a person's relative strengths and weaknesses (i.e., they are helpful in conducting an intra-individual assessment of phonological processing assets and deficits). Examiners can not only analyze performance across the contents of the subtests but also investigate a person's status relative to the constructs incorporated into the test (i.e., composite scores). By considering the CTOPP-2's composite scores, examiners can contrast aspects pertaining to phonological awareness, phonological memory, and rapid naming abilities.

Third, examiners can use the CTOPP-2 to provide a means for evaluating an individual's progress in prescribed remedial programs. The monitoring of an individual's progress is an important component of special instruction and is often required by law or by school policy. Periodic assessment is desirable even when it is not required because it provides educators with evidence that the instructional program is appropriate to meet an individual's needs.

Fourth, the CTOPP-2 has value as a research tool, especially for investigators who wish to study phonological processing using standardized instruments. Its results can be used to test the adequacy of various theories of phonological processing; to measure the relationship of certain abilities to reading proficiency,

to successful everyday living skills, and so forth; and to determine the effectiveness of various intervention programs on test performance.

Tests that are used for these four purposes should satisfy rigorous technical adequacy criteria. In fact, many state agencies and several professional organizations, most notably the American Psychological Association (1985), have set up criteria that all good tests should satisfy. These standards generally relate to reliability, validity, normative data, and methods for reporting scores. The CTOPP-2 was constructed with these requirements clearly in mind in order to better meet the purposes for testing.

3 Recording and Interpreting the CTOPP-2 Results

This chapter explains how to record and interpret the CTOPP-2 scores. Specifically discussed are (a) the Examiner Record Booklet, (b) the normative scores and their interpretation, (c) the use of intra-test discrepancy analyses, and (d) cautions that must be considered when interpreting test scores.

Completing the Examiner Record Booklet

The Examiner Record Booklet is used to record and summarize the examinee's performance. This booklet contains five sections: Identifying Information, Subtest Performance, Composite Performance, Descriptive Terms, and Administration and Scoring Instructions for Subtests. An example of the first page of Joshua's completed Examiner Record Booklet is provided in Figure 3.1.

Section 1. Identifying Information

In Section 1, space is provided to record the examinee's name, gender, grade, and school information; to calculate the child's age; and to record the examiner's name and title. Calculate the examinee's age at the time of testing by subtracting his or her birth date from the date on which the testing took place. For example, as shown in Figure 3.1, Joshua was born on June 4, 2004, and was tested on March 1, 2012. Because 4 cannot be subtracted from 1, 30 days (i.e., 1 month) are borrowed from the adjacent month's column and added to the 1 day. The date of testing is now 2012-2-31. Because 6 months cannot be subtracted from 2 months, 12 months (i.e., 1 year) are borrowed from the adjacent year's column. The date of testing becomes 2011-14-31. Simple subtraction is applied, and Joshua's age is found to be 7 years, 8 months, and 27 days. For purposes of using the normative tables, **do not round ages upward**. Thus, Joshua is 7 years 8 months of age, not 7 years 9 months.

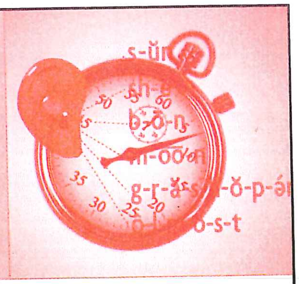
Section 2. Subtest Performance

In Section 2, the examiner records the examinee's raw score, age and grade equivalents, percentile rank, scaled score, and descriptive term for each subtest. The raw scores are recorded first. Age and grade equivalents that correspond to the raw scores are found in Appendix A and are recorded next. These are followed by the percentile ranks and scaled scores, which are located in the normative tables in Appendix B. Complete descriptions of age and grade equivalents, percentile ranks, and scaled scores are provided later in this chapter.

Comprehensive Test of Phonological Processing–Second Edition

CTOPP-2

Examiner Record Booklet Ages 7–24



Richard K. Wagner Joseph K. Torgesen Carol A. Rashotte Nils A. Pearson

Section 1. Identifying Information

Name Joshua Jones

Female ☐ Male ☒ Grade 2

Parent/Guardian

School Covert Avenue School

	Year	Month	Day
Date Tested	~~2012~~ 2011	~~3~~ ~~2~~ 14	~~1~~ 31
Date of Birth	2004	6	4
Age*	7	8	27

*When accessing the normative tables, use years and months. Do not round up.

Teacher's Name Mr. Rouner

Examiner's Name Dr. Joseph Craffer

Examiner's Title Psychologist

Section 2. Subtest Performance

Subtest	Raw Score	Age Equiv.	Grade Equiv.	%ile Rank	Scaled Score	*SEM*	Descriptive Term
Core							
1. Elision (EL)	17	6-9	1.7	25	8	1	Avg.
2. Blending Words (BW)	15	5-9	K.7	9	6	1	Below Avg.
3. Phoneme Isolation (PI)	16	6-3	1.2	16	7	1	Below Avg.
4. Memory for Digits (MD)	12	4-6	LK	9	6	1	Below Avg.
5. Nonword Repitition (NR)	14	6-0	1.0	25	8	1	Avg.
6. Rapid Digit Naming (RD)	38	5-9	K.7	9	6	1	Below Avg.
7. Rapid Letter Naming (RL)	40	6-0	1.0	16	7	1	Below Avg.
Supplemental							
8. Blending Nonwords (BN)	15	6-9	1.7	25	8	1	Avg.
9. Segmenting Nonwords (SN)	14	6-3	1.2	25	8	1	Avg.

Section 3. Composite Performance

Composite	EL	BW	PI	MD	NR	RD	RL	BN	SN	Sum of Scaled Scores	%ile Rank	*SEM*	Composite Score	Descriptive Term
	Subtest Scaled Score													
Phonological Awareness	8	6	7							21	12	4	82	Below Avg.
Phonological Memory				6	8					14	12	6	82	Below Avg.
Rapid Symbolic Naming						6	7			13	8	4	79	Poor
Alt. Phonological Awareness								8	8	16	21	4	88	Below Avg.

Section 4. Descriptive Terms

Scaled Score	1–3	4–5	6–7	8–12	13–14	15–16	17–20
Descriptive Term	Very Poor	Poor	Below Average	Average	Above Average	Superior	Very Superior
Composite Score	<70	70–79	80–89	90–110	111–120	121–130	>130

Figure 3.1. Sample page 1 of Examiner Record Booklet filled out for Joshua.

For example, Joshua scored 17 points on the Elision subtest. This score converts to an age equivalent of 6-9 and a grade equivalent of 1.7 (see Table A.1 in Appendix A). For the purposes of using the proper normative table, his age was not rounded up and is 7-8. The examiner consults Table B.9 in Appendix B to transform his raw score to a percentile or standard score. According to this table, a raw score of 17 is transformed into a percentile rank of 25 and a scaled score of 8. Descriptive terms are found in Table 3.1.

Section 3. Composite Performance

Scaled scores for the subtests are recorded in the spaces indicated for them in this section. The scaled score for each subtest is assigned to the Phonological Awareness, Phonological Memory, Rapid Symbolic Naming, Alternate Phonological Awareness, or Rapid Non-Symbolic Naming composite.

On the right side of Section 2 on the Examiner Record Booklet, the standard scores for the subtests are assigned to the composites that represent constructs in the model that was used to build the test. For example, the standard scores for Elision (EL), Blending Words (BW), and Phoneme Isolation (PI) are summed, and the summed value is converted into a Phonological Awareness Composite Score (PACS). For children ages 4–6 years, the standard scores for Elision (EL), Blending Words (BW), and Sound Matching (SM) are summed, and the summed value is converted into a PACS using Appendix C. The remaining composites are computed by combining the standard scores that comprise them.

For example, to calculate the PACS for Joshua, the standard scores for the EL, BW, and PI subtests are summed and transformed into the PACS by consulting the table in Appendix C. In Joshua's case, the sum of the standard scores (8 + 6 + 7) is 21, which translates into a PACS of 82.

The scaled scores that make up each composite are summed and recorded in the Sum of Scaled Scores oval. This summed value is converted into the Phonological Awareness Composite Score, the Phonological Memory Composite Score, the Rapid Symbolic Naming Composite Score, the Alternate Phonological Awareness Composite Score, and the Rapid Non-Symbolic Naming Composite Score, and a percentile rank using Table C.1 in Appendix C. These composite scores and percentile ranks are recorded in their appropriate spaces at the right side of Section 3.

For example, Joshua received the following scaled scores for the nine subtests: Elision, 8: Blending Words, 6; Phoneme Isolation, 7; Memory for Digits, 6; Nonword Repetition, 8; Rapid Digit Naming, 6; Rapid Letter Naming, 7; Blending Nonwords, 8; Segmenting Nonwords, 8. To calculate the CTOPP-2 Phonological Awareness Composite Score, the scaled scores for the Elision, Blending Words, and Phoneme Isolation subtests are summed. This sum can be transformed into a composite score or percentile rank by consulting Table C.1. In Joshua's case, the sum of the three subtests is 21. The examiner locates the column sum of 3 subtests and reads down the column until 21 is found. The corresponding composite score is 82. The percentile rank is 12. An composite score of 82 indicates that Joshua's phonological awareness performance is below average (see Section 4). The remaining composites are completed in the same manner.

Table 3.1
Descriptive Terms for CTOPP-2 Scaled Scores and Composite Scores

Scaled score	Composite score	Descriptive term
17–20	>130	Very superior
15–16	121–130	Superior
13–14	111–120	Above average
8–12	90–110	Average
6–7	80–89	Below average
4–5	70–79	Poor
1–3	<70	Very poor

Section 4. Descriptive Terms

Descriptive terms that correspond to scaled scores and composite scores are provided in this section. The terms range from *very poor* to *very superior.*

Section 5. Administration and Scoring Instructions for Subtests

Instructions for giving and scoring the CTOPP-2 subtests are provided at the beginning of each subtest in Section 5 of the Examiner Record Booklet.

Test Scores and Their Interpretation

The CTOPP-2 yields six types of normative scores: age and grade equivalents, percentile ranks, subtest scaled scores, composite scores, and developmental scores. Information about how these scores are derived and how to interpret them is discussed in this section.

Age and Grade Equivalents

Age and grade equivalents for tests of specific abilities are usually labeled according to the content of the test. Thus, age and grade equivalents associated with tests of reading are called *reading ages/grades,* and those associated with tests of visual perception are called *phonological processing ages/grades.* These scores are derived by calculating the average normative group score at each 6-month interval. Through the process of interpolation, extrapolation, and smoothing, age equivalents are generated for each raw-score point achieved on a subtest. Thus, a raw score of 9 on the Elision subtest yields an age equivalent of 5 years 6 months and a grade equivalent of K.4 (see Table A.1 in Appendix A).

The use of age and grade equivalents has come under close scrutiny in recent years, so much so that the American Psychological Association and most other authorities advise strongly against their use on the grounds that their statistical properties are inadequate and their results are often misleading. However, today the use of age and grade equivalents are mandated. Because these scores are frequently required for administrative purposes, we provide them (reluctantly). Table A.1 in Appendix A is used to convert subtest raw scores to age and grade equivalents.

Because age and grade equivalents are problematic, we recommend that CTOPP-2 users read the cautions associated with age equivalents found in the works of Aiken and Groth-Marnat (2006), Anastasi and Urbina (1997), Linn and Miller (2005), and Salvia, Ysseldyke, and Bolt (2010). We prefer to use scaled scores, composite scores, and percentile ranks rather than age equivalents when reporting results to parents and other professionals.

Percentile Ranks

Percentile ranks represent values that indicate the percentage of the distribution that is equal to or below a particular score. For example, a percentile rank of 56 means that 56% of individuals in the same age range of the standardization sample scored at or below the examinee's score. Because this interpretation is easy to understand, percentiles are popular scores for practitioners to use when sharing test results with others. Note that the distance between two percentile ranks becomes much greater as those ranks are more distant from the mean or average (i.e., the 50th percentile). Percentile ranks are generated for the subtests and composites using tables in Appendixes B and C, respectively.

Subtest Scaled Scores

Scaled scores are a type of standard score that provides the clearest indication of a child's subtest performance. Based on a normal distribution with a mean of 10 and standard deviation of 3, subtest standard scores are converted from raw scores using Tables B.1 through B.17 in Appendix B. Guidelines for interpreting scaled scores (and composite scores, discussed next) are in Table 3.1.

If done wisely, the interpretation of subtest performance will yield some information about a person's strengths and weaknesses. However, examiners should not place too much reliance on the interpretation of subtest results. For example, the reliability associated with the CTOPP-2 subtests is not as high as that of the CTOPP-2 composites. This means that interpretations, diagnoses, and judgments made on the basis of subtest scores are going to contain more error than those based on composite scores.

Evaluation of subtest performance is useful in generating hypotheses or speculations about why a person did well or poorly on a composite, but important decisions about diagnosis and placement should rest primarily on the interpretation of the composite scores. When interpreting subtests, do not generalize beyond the guidelines listed below.

Elision measures the ability to remove phonological segments from spoken words to form other words.

Blending Words measures the ability to synthesize sounds to form words.

Sound Matching measures the ability to select words with the same initial and final sounds.

Phoneme Isolation measures the ability to isolate individual sounds within words.

Memory for Digits measures the ability to repeat numbers accurately.

Nonword Repetition measures the ability to repeat nonwords accurately.

Rapid Digit Naming measures the ability to rapidly name digits.

Rapid Letter Naming measures the ability to rapidly name letters.

Rapid Color Naming measures the ability to rapidly name colors.

Rapid Object Naming measures the ability to rapidly name objects.

Blending Nonwords measures the ability to synthesize sounds to form nonwords.

Segmenting Nonwords measures the ability to segment nonwords into phonemes.

Composite Scores

Composite scores are the most useful scores on the CTOPP-2 in part because they are the most reliable. They are another type of standard score and are calculated by applying a direct linear transformation to the sum of scaled scores to obtain a distribution with a mean of 100 and a standard deviation of 15. The resulting data across age levels were smoothed somewhat to allow for consistent progression. The sums of scaled scores for the composites are easily converted into composite scores using Table C.1, found in Appendix C. Guidelines for interpreting composite scores are given in Table 3.1. The composite scores discussed in this section relate to the theoretical model underlying the CTOPP-2, which was described in Chapter 1. We pointed out that phonological processes could be conceptualized as phonological awareness, phonological memory, and rapid naming. The CTOPP-2 subtests were combined in such a way as to form composites to represent these three constructs.

What the Composites Measure.

Three composite scores are generated that reflect on the examinee's status relative to the phonological constructs that are incorporated into the CTOPP-2. These composite scores indicate an examinee's ability relative to phonological awareness, phonological memory, and rapid naming.

Phonological Awareness Composite Score.
The Phonological Awareness Composite Score comprises the standard scores of three subtests: Elision, Blending Words, and Sound Matching for 4–6-year-olds

and Elision, Blending Words, and Phoneme Isolation for 7–24-year-olds. The PACS measures an individual's phonological awareness—that is, awareness of and access to the phonological structure of oral language. As discussed in Chapter 1, this awareness proceeds from (a) word-length units in compound words, to (b) syllables within words, to (c) onset-rime units within syllables, to (d) individual phonemes within rimes, and finally to (e) individual phonemes within consonant clusters. Similarly, awareness proceeds from (a) sensitivity to same versus different phonological segments, to (b) an ability to identify and count phonological segments, to (c) an ability to manipulate phonological segments. Because the phonological structure of oral language is represented in written language—for alphabetic writing systems, letters roughly correspond to phonemes—phonological awareness provides a beginning reader with an important tool for understanding relations between written and spoken language.

Children with well-developed phonological awareness learn to read more easily than do children with poorly developed phonological awareness. A deficit in phonological awareness is viewed as the hallmark of reading disability or dyslexia. Poor phonological awareness is associated with poor reading for both individuals whose poor reading levels are discrepant from their IQs and individuals whose poor reading levels are consistent with their IQs.

Of the three kinds of phonological processing measured by the CTOPP-2, phonological awareness appears to be the most responsive to interventions. For individuals with a deficit in phonological awareness who are at or below the initial stage of reading acquisition, intervention designed to promote phonological awareness may prove fruitful. For individuals with a deficit in phonological awareness who are beyond the initial stage of reading acquisition, a more productive way to enhance phonological awareness is likely to be in the context of a reading program that is structured and systematic and that explicitly points out connections between spoken and written language (e.g., letter–sound correspondences, blending skills).

Phonological Memory Composite Score.
The Phonological Memory Composite Score comprises the standard scores of two subtests—Memory for Digits and Nonword Repetition—for all individuals. The PMCS represents the examinee's ability to code information phonologically for temporary storage in working or short-term memory. Specifically, the PMCS provides an assessment of the functioning of the part of memory called the *phonological loop*, which provides a brief, verbatim storage of auditory information. The phonological loop itself comprises (a) a phonological store that records the most recent 2 seconds' worth of auditory information and (b) an articulatory control process that provides input to the phonological loop initially and also can refresh information already in the loop so that it can be stored for longer than 2 seconds.

A deficit does not inevitably lead to poor reading of familiar material but is more likely to impair decoding of new words, particularly words that are long enough to decode bit by bit, as a means of storing intermediate sounds. A deficit in phonological memory may not impair listening or reading comprehension for simple sentences but is likely to impair both listening and reading comprehension for more complex sentences.

Rapid Symbolic Naming Composite Score.
The Rapid Symbolic Naming Composite Score comprises the standard scores of two subtests—Rapid Digit Naming and Rapid Letter Naming. The abilities measured by the RSNCS include efficient retrieval of phonological information from long-term, or permanent, memory and quick and repeated execution of a sequence of operations. Efficient retrieval of phonological information and execution of sequences of operations are required when readers attempt to decode unfamiliar words. Because of the timed nature of the subtests that make up the RSNCS, individuals who perform poorly commonly have problems with reading fluency. Individuals who have deficits in both rapid naming and phonological awareness appear to be at greater risk of reading problems compared to individuals with deficits in only one of the two areas.

Rapid Non-Symbolic Naming Composite Score.
The RNNCS comprises the standard scores of two subtests—Rapid Color Naming and Rapid Object Naming—and offers an alternative for young children not familiar with letters and numbers.

The abilities measured by the RNNCS include efficiently retrieving phonological information from long-term, or permanent, memory and executing a sequence of operations quickly and repeatedly using objects and colors.

Alternate Phonological Awareness Composite Score.
As previously described in Chapter 1, an alternate composite is available for 7–24-year-olds, consisting of Blending Nonwords and Segmenting Nonwords. This Alternate Phonological Awareness Composite (APACS) is available for examiners who desire to assess phonological awareness exclusively with nonwords. If done wisely, the interpretation of subtest performance will yield much useful information about a person's strengths and weaknesses. Still, there are problems regarding subtests that must be faced squarely when conducting such an analysis. For example, the reliability associated with subtests is generally lower than that associated with composites. This means that interpretations, diagnoses, and judgments made on the basis of subtest scores are going to contain considerably more error than those based on composite scores. This is an observation that is true of all tests.

Developmental Scores

Developmental scores provide an equal-interval developmental scale that is useful for examining changes in performance over time. An equal-interval scale is one for which the difference in underlying ability represented by scores of, say, 450 and 500 is the same as the difference represented by scores of 500 and 550. Developmental scores are used by researchers in longitudinal studies, but they are useful for practitioners who want to examine changes in absolute levels of performance over time or in response to intervention. Unlike relative measures of performance, such as percentile ranks or standard scores, that are based on a comparison between an individual's score and the scores of age-matched peers in the normative sample, absolute measures of performance such as developmental scores place an individual on a scale that represents the skill level measured by the test items. In this sense, developmental scores are more like

criterion-referenced scores than norm-referenced scores. They can be thought of as raw scores that have been rescaled on an equal-interval developmental scale.

Developmental scores are available for all non-timed CTOPP-2 subtests (Elision, Blending Words, Sound Matching, Phoneme Isolation, Memory for Digits, Nonword Repetition, Blending Nonwords, and Segmenting Nonwords) and composites (Phonological Awareness, Alternate Phonological Awareness, and Phonological Memory). Developmental scores for composites are calculated by averaging the developmental scores of the subtests that are included in the composite. Developmental scores are not available for the timed rapid naming subtests (Rapid Naming of Digits, Rapid Naming of Letters, Rapid Naming of Objects, and Rapid Naming of Colors) because the item response theory (IRT) model used is not appropriate for timed measures.

The IRT modeling used both for test construction and for calculating developmental scores is described in Chapter 6. Initially, a one-parameter Rasch model and either a two- or three-parameter (depending on whether the response format was free response or multiple choice) model were fit for each subtest using the data provided by the normative sample. Because the correlation between parameter estimates derived from the Rasch model and the two- or three-parameter models approached the reliabilities of the subtests, the Rasch model was used to calculate developmental scores. An advantage of using the Rasch model to calculate developmental scores is that an examiner can convert raw scores to developmental scores using a simple conversion table. For the CTOPP-2, Table D.1 in Appendix D is a conversion table for converting raw scores for CTOPP-2 subtests to developmental scores. To calculate developmental scores for composites, average together the developmental scores of the subtests included in the composite.

Conducting Discrepancy Analyses

Examiners will need to know when the difference (called a *difference score*) between the PACS, PMCS, RNNCS, and RSNCS is large enough to be noteworthy. Two methods can be used to examine the importance of difference scores. The first method indicates whether the difference score is *statistically* significant; the second indicates whether it is large enough to be considered *clinically* significant. Generally, researchers are interested in the former; teachers and clinicians are interested in the latter.

Anastasi and Urbina (1997) provided a formula (p. 111) for determining how large a difference score must be to be statistically significant. We used that formula to compile the minimal difference scores required for significances (at the $p < .05$ level) for the two composite scores. A statistically significant difference does not necessarily mean that the difference is large enough to be clinically useful (Kaufman & Kaufman, 1990). The use of statistical significance alone will identify too many false-positive cases. Reynolds (2003) provided a formula (#7, p. 483) to determine how large the difference score must be to be considered clinically useful.

The formulas referred to in the previous paragraph were used to compile difference scores for both statistical and clinical usefulness. Difference scores